

This is an author-generated copy of the paper published in: Genome. 49:209-218
(2006) available online from: http://pubs.nrc-cnrc.gc.ca/cgi-bin/rp/rp2_desc_e?gen

**Molecular Structure and Organization of the Wheat Genomic Manganese Superoxide
Dismutase Gene**

Kwang-Hyun Baek, Daniel Z. Skinner, Peng Ling, and Xianming Chen

Kwang-Hyun Baek

Department of Crop and Soil Sciences, 210 Johnson Hall, Washington State University, Pullman,
WA 99164-6420, USA

Daniel Z. Skinner¹

USDA/ARS, Department of Crop and Soil Sciences, 209 Johnson Hall, Washington State
University, Pullman, WA 99164-6420, USA

Peng Ling

USDA/ARS, Department of Plant Pathology, 209 Johnson Hall, Washington State University,
Pullman, WA 99164-6420, USA

Xianming Chen

USDA/ARS, Department of Plant Pathology, 360 Johnson Hall, Washington State University,
Pullman, WA 99164-6420, USA

¹Corresponding author (email: dzs@wsu.edu, Fax: 509-335-2553)

Abstract

The genomic structure of an MnSOD gene in wheat was elucidated by sequencing a clone from a BAC library of a stripe-rust resistant wheat line. The clone was identified by hybridization with a wheat MnSOD cDNA. The gene consisted of six exons interrupted by five introns with a total length of 4770 nucleotides from the start codon to the termination codon. The wheat MnSOD gene was the longest among those sequenced from plant species. The transcription initiation site was preceded by a G+C-rich promoter without a TATA or CAAT box. The promoter contained many putative cis-acting regulatory elements including an ABA-responsive element, stress-responsive element, and GC-repeat, and several other structural features in common with the promoter of the rice MnSOD gene. A *Stowaway*-like transposable element was found in intron 5 of the wheat MnSOD gene, but further investigation revealed the transposable element was not present in all copies of the MnSOD genes

Keywords: AY963808, BAC clone, manganese superoxide dismutase, gene, promoter, transposable element.

Abbreviations: MITEs: miniature inverted-repeat transposable elements,,

MnSOD: manganese superoxide dismutase, ROS: reactive oxygen species,

SOD: superoxide dismutase

Introduction

Reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2), superoxide anion (O_2^-), and the hydroxyl radical (OH^\cdot) can be detrimental to living organisms because of their high oxidizing potentials in living cells (Bowler et al. 1992). ROS are generated on a regular basis in biological pathways as by-products or as signal transducers (Wu et al. 1997; Zelko et al. 2002). However, excessive production or inefficient scavenging of ROS can cause over-accumulation which can injure or kill cells (Monk et al. 1989; Fridovich 1991). All basic molecules in living organisms can be attacked by ROS, e.g., lipids, carbohydrates, proteins, and nucleic acids (Kendall and McKersie 1989, Sato et al. 1993; Fucci et al. 1983; Halliwell and Gutteridge 1999).

Antioxidant enzymes in living organisms have evolved as very sophisticated and effective scavengers of ROS (Allen 1995; Halliwell and Gutteridge 1999).

Superoxide dismutase (SOD; superoxide: superoxide oxidoreductase, EC. 1.15.1.1) is an essential antioxidant enzyme protecting many cellular components by converting two superoxide anions (O_2^-) into H_2O_2 and O_2 (Fridovich 1991; Fink and Scandalios 2002). SODs are found in many organisms including all oxygen-consuming organisms, aero-tolerant anaerobes, and some obligate anaerobes (Fink and Scandalios 2002). There are several isozymes of SOD classified by the location and the catalytic metals required (Halliwell and Gutteridge 1999). Manganese superoxide dismutase (MnSOD) found in mitochondria or peroxisomes has Mn (III) at the active site. Iron superoxide dismutase requires Fe (III) at the active site and is found in chloroplasts. Copper, zinc superoxide dismutase needs two catalytic metals, Cu(II) and Zn(II), at the active site and is located both in chloroplasts and cytosol. Nickel superoxide dismutase has a Ni (II/III)

pair at the active site and is found in the *Streptomyces* genus (Fink and Scandalios 2002).

MnSOD is the primary protective enzyme for oxidative stress in mitochondria (Bowler et al 1991; Møller 2001) and has been proven to be the only form of SOD essential for the survival of aerobic life (Carlioz and Touati 1986). MnSOD genes from phylogenetically distant organisms have striking homology at the amino acid and nucleotide level (Fink and Scandalios 2002).

MnSOD precursor protein is encoded in the nucleus, is targeted to the mitochondrial matrix, and is processed to the active MnSOD enzyme form after removal of the mitochondria-targeting leader sequence (Bowler et al. 1989). The MnSOD holoenzyme in plants is composed of four subunits with a total molecular mass of around 91 kDa (Streller et al. 1994). The role of MnSOD in plants has been extensively studied due to the unique property of mitochondrial protection and a possible role in tolerance of environmental stresses, such as chilling, freezing, oxidative stress, aluminum toxicity, etc. There are many specific examples of an advantage of high expression of MnSOD, such as enhanced tolerance to: chilling stress in *Chlorella* with a high amount of MnSOD (Clare et al. 1984), chilling temperatures in transgenic *Zea mays* expressing a tobacco MnSOD gene (Breusegem et al. 1999), freezing stress in transgenic *Medicago sativa* expressing a tobacco MnSOD gene (McKersie et al. 1993), oxidative stress in tobacco with overproduction of MnSOD (Yu et al. 1999), and oxidative stress and aluminum toxicity in transgenic *Brassica* plants expressing wheat MnSOD (Basu et al. 2001).

In wheat ($2n=6x=42$), the expression of total MnSOD transcripts increased significantly in both spring and winter wheat seedlings in response to cold (Wu et al. 1999; Baek and Skinner, 2003). MnSOD genes in wheat have been mapped to the long arms of the homeologous group-2

chromosomes (Wu et al. 1999), and mapping of 58 wheat ESTs revealed a total of five loci on the long arms of the 2A, 2B, and 2D chromosomes (<http://wheat.pw.usda.gov/GG2/index.shtml>), indicating at least five MnSOD genes are present in the hexaploid genome. MnSOD genes in wheat were further identified as a multigene family based on alignment of the sequences of wheat MnSOD cDNAs in GenBank (accessions AF092524, U72212, and U73172) and of cDNA clones isolated in our laboratory, genetic mapping data (Wu et al. 1999), and EST sequence and mapping data (<http://wheat.pw.usda.gov/GG2/index.shtml>). The nucleotide and amino acid sequence of each MnSOD member in the multigene family are slightly different though the final gene products, MnSOD enzyme, have the same enzymatic function.

The genomic DNA of MnSOD genes have been cloned from *Arabidopsis thaliana* (GenBank accession GI 12408720), *Hevea brasiliensis* (GenBank accession GI 348136), *Pisum sativa* (GenBank accession GI 945043), and *Oryza sativa* (GenBank accession GI 6440979). The genomic structure of the *Hevea brasiliensis* MnSOD gene, including 5' and 3' regions, was analyzed by Miao and Gaynor (1993). The *Hevea* MnSOD gene had a putative TATA box (ATATAT) located thirty nucleotides upstream of the transcription start site and also had two tandem 'TGACG'-like boxes from position -87 to -66 upstream of the transcription start site believed to be involved in binding of transcription factor ASF-1 for initiating gene expression in roots (Miao and Gaynor 1993). All MnSOD genes from plant species studied so far share the same basic structural format, six exons and five introns (Fink and Scandalios 2002). Furthermore, the lengths of the exons were very similar in *Arabidopsis*, *Hevea*, *Pisum*, and *Oryza*.

There is increasing interest in the regulation of MnSOD genes in plants for enhancing tolerance

of environmental stresses; however, there is insufficient information available on the gene structure and regulatory elements of the wheat MnSOD gene. Therefore, our objective was to clone a wheat genomic MnSOD gene, and to analyze the physical structure to gain an understanding of the possible mechanisms underlying gene regulation.

Materials and Methods

BAC Library Construction and Screening

High molecular weight DNA was extracted from the *Yr5* near-isogenic line BC₆, developed by backcrossing the *Yr5* donor '*Triticum spelta* album' with the recurrent parent 'Avocet' (Wellings et al 2004). *Yr5* near-isogenic line BC₆ is resistant to almost all races tested of *Puccinia striiformis* Westend. f. sp. *Tritici* Eriks., the pathogen inciting stripe rust. The DNA was partially digested with *Hind*III and cloned into pECBAC1 vector using electroporation of DH10B *E. coli*. The transformed *E. coli* was grown on LB agar containing 12.5 µg/ml of chloramphenicol. Colonies potentially carrying wheat DNA inserts were identified by conventional β-galactosidase color selection. BAC clones were arrayed into a total of 1,100 384-well micro-plates. The average size of the BAC clone inserts was 130 kb. A complete description of the BAC library will appear elsewhere (Ling and Chen, 2006).

An estimated 1.6 X hexaploid wheat genome coverage (230,400 clones) was spotted onto 22.5 cm x 22.5 cm Hybond N+ nylon membrane (Amersham Pharmacia Biotech, UK) with 4 x 6 x 4 formation using a Q array robot system (Genetix, New Milton, UK). The nylon membranes were placed on plates of LB with 1.5 % agar containing 12.5 µg/ml of chloramphenicol, and the *E. coli* was grown at 37 °C for 22 hr. The membranes were then treated with 10 % SDS for 6 min,

then denaturing solution (0.5N NaOH/ 1.5 M NaCl) for 10 min, then were dried at room temperature for 20 min. Membranes were then neutralized in 1 X neutralizing solution (0.5 M Tris-Cl/1.5 M NaCl) for 5 min, 0.1 X neutralizing solution for 5 min, 2 X SSC/ 0.1 % (w/v) SDS for 5 min, 2X SSC for 5 min, and 50 mM Tris-HCl (pH 7.5) for 5 min. After being allowed to air-dry, the nylon membranes were baked at 80 °C for two hours.

The membranes were then probed with cDNA of wheat MnSOD to screen the BAC clones for the MnSOD gene. The probe consisted of cDNA of the wheat MnSOD gene from the start codon to the termination codon (706 bp), or from the start codon to the middle of the gene (346 bp). The cDNA was generated and amplified using a Superscript one-step kit (Invitrogen, San Diego, USA) using primers forward: 5' - CACACACCAAACACACTATCCATG and reverse: 5' - TCATGCAAGCACTTTCTCATACTCT (for 706 bp) or 5' - AGGCTTGAGGTTCTTCCAGA (for 346 bp). The RT-PCR profile consisted of 15 min cDNA synthesis at 50 °C followed by 3 min denaturation at 95°C, then 32 cycles of 30 s at 95°C, 40 sec at 54 °C, and 1 min at 72°C; then 5 min at 72°C. The RT-PCR solution was composed of 1 X reaction buffer and 0.4 µl reverse transcriptase/Taq polymerase mix from the Superscript one-step RT-PCR kit, 2 mM MgCl₂, 100 nM primers, and 125 ng total RNA in 20 µl reactions covered with 15 µl mineral oil. PCR products were cloned with Topo-XL PCR cloning kit (Invitrogen, San Diego, USA). Plasmids from white positive clones were sequenced to confirm the clones were of wheat MnSOD at the Sequencing Center of Washington State University, Pullman.

Labeling was performed by using a Random Primer Labeling Kit (Gibco-BRL, UK) and ³²P. The nylon membranes were prehybridized at 65 °C for 1 hr in 5 ml of hybridization buffer (pH 7 0.5

M sodium phosphate buffer, 7 % SDS, and 1 mM EDTA). The heat-denatured probe was added and hybridization was carried out at 65 °C for 16 hr. The membranes were then washed at 25 °C in 2X SSC/ 0.1 % SDS (w/v), 1 X SSC/ 0.1 % SDS (w/v) for 5 min each, and at 65 °C in 0.1 X SSC/ 0.1 % SDS (w/v) for 10 min. Autoradiography was performed with Kodak X-Omat Blue 35.6 X 43.2 cm film (NEN Life Sciences, Boston, MA) with the aid of intensifying screens at -70 °C for 2 days.

Subcloning of the Positive BAC Clones

Plasmids were extracted from the BAC clones showing strong hybridization to the wheat MnSOD cDNA probes by alkaline lysis (Sambrook et al. 1989), then were digested with *Xba*I and analyzed on 1 % agarose gel in 1 x modified TAE (pH 8.0, 40 mM Tris-acetate, 0.1 mM Na₂EDTA). The digested plasmid fragments were transferred to a nylon membrane, and the nylon membrane was probed with ³²P-labeled MnSOD cDNA as described above. Five BAC clones identified as nos. 6, 18, 55, 56, and 59 had strong hybridization signals of 7-9 kb fragments. The fragments were isolated from 1 % agarose gels and cloned into the *Xba*I site of pBlueScript SK(+) plasmid (Stratagene, La Jolla, CA) in Top 10 *E. coli* strain (Invitrogen, San Diego, CA) using electroporation. The transformed *E. coli* was grown on 1.5 % agarose gels, transferred to a nylon membrane, and lysis and screening was carried out as above, again with MnSOD cDNA as the probe. PCR also was used to confirm the presence of MnSOD genes in these subclones. Based on conserved exon sequences from published plant MnSOD genes, primers were designed to flank intron 5. The primers used were, forward: 5'-TCCTTTGTTGGGAATTGATGTC and reverse: 5'-TCATGCAAGCACTTTCTCATACTCT. The PCR profile was 3 min denaturation at 94°C, then

32 cycles of 30 s at 94°C, 40 sec at 52 °C, and 1.5 min at 72 °C; then, 5 min at 72°C. The PCR solution was composed of 1 X reaction buffer and 2 unit of *Taq* polymerase from Promega (Madison, WI, USA), 2.2 mM MgCl₂, 100 nM primers, and 500 ng total DNA in 25 µl reaction solution covered with 20 µl mineral oil.

Partial sequencing of the subcloned *Xba*I fragments that again hybridized to MnSOD cDNA and yielded a PCR product using the MnSOD based primers suggested that BAC clone number 56 contained a complete MnSOD gene. Southern analysis of *Hind*III-digested clone 56 indicated most of the MnSOD gene was on an 8.5 kb *Hind*III fragment. This fragment was subcloned as above. The 8.5 kb *Hind*III fragment, a 2.1 kb and a 7 kb *Xba*I fragment were sequenced using transposon-mediated sequencing. The EZ:: TNTM <Kan-2> insertion kit (Epicentre Technologies, Madison, WI) was used and subclones were sequenced to approximately 3X coverage. The resulting partial sequences and the sequence of a PCR fragment containing the 3' UTR (primers, forward: 5'- TCCTTTGTTGGGAATTGATGTC and reverse:

5'-AACAGCACTAGCGAACGAGTT) were assembled into the complete sequence (11,304 bp) using CAP3 program (<http://deepc2.zool.iastate.edu/aat/cap/cap.html>; Huang and Madan 1999).

The software used to investigate structural features were: DNA sequence alignment: Multalign, <http://prodes.toulouse.inra.fr/multalin/multalin.html> (Corpet, 1988); DNA sequence comparison:

BLAST, <http://www.ncbi.nlm.nih.gov/BLAST> (Altschul *et al.*, 1997; Tatusova and

Madden,1999); Exon-Intron prediction: Splice View,

<http://l25.itba.mi.cnr.it/~webgene/wwwspliceview.html> (Rogozin and Milanesi, 1997); Gene

prediction: FGENESH, <http://sun1.softberry.com/berry.phtml?topic=fgenes&group=>

[programs&subgroup=gfind](http://sun1.softberry.com/berry.phtml?topic=fgenes&group=programs&subgroup=gfind) (Salamov and Solovyev, 2000), and GeneScan,

<http://genes.mit.edu/GENSCAN.html> (Burge and Karlin, 1997); Transcription start site prediction: TRANSFAC, http://sdmc.lit.org.sg/promoter/dragonGSF1_0/genestart.htm (Matys et al., 2003); CpG island prediction: CpG Island Searcher, <http://cpgislands.usc.edu/cpg.aspx>, (Takai and Jones, 2002; Takai and Jones, 2003); Plant promoters prediction: TSSP, <http://www.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter>; Cis-acting regulatory elements prediction: PlantCARE, <http://intra.psb.ugent.be:8080/PlantCARE> (Lescot et al., 2002); Polyadenylation signal: HCPolya, http://l25.itba.mi.cnr.it/~webgene/wwwHC_polya.html (Milanesi et al., 1996); and RNA secondary structure: Mfold, <http://www.bioinfo.rpi.edu/applications/mfold/old/rna/form1.cgi> (Mathews et al., 1999; Zuker, 2003).

Cloning of a Stowaway-like transposable element embedded in intron 5

The primers designed to flank intron 5 also were used to amplify a *Stowaway*-like element found in that intron. The forward primer: 5'- TCCTTTGTTGGGAATTGATGTC in exon 5 and the reverse primer: 5'- TCACGCAAGCACTTTTTCATACTCT in exon 6 were used. The PCR profile was 3 min denaturation at 94°C, then 32 cycles of 30 s at 94°C, 40 sec at 52 or 58 °C annealing temperature, and 1.5 min at 72°C; then, 5 min at 72°C. The PCR solution was composed of 1 X reaction buffer and 2 unit of *Taq* polymerase from Promega (Madison, WI, USA), 2.2 mM MgCl₂, 100 nM primers, and 500 ng total DNA in 25 µl reaction solution covered with 20 µl mineral oil. The PCR fragments were purified with a DNA gel extraction kit (Millipore corporation, Bedford, MA), cloned into pCR® II vector (TA Cloning® Kit, Invitrogen, San Diego, CA), and sequenced at the DNA sequencing facility at Washington State University, Pullman.

Results

Sequencing and Analysis of the MnSOD gene

The complete sequence consisted of 11,304 nucleotides (GenBank accession AY963808). The MnSOD gene comprised bases 6,330 (translation start) to 11,304 (the end of 3' UTR) (Fig. 1). The wheat MnSOD gene consisted of 4770 nucleotides from the start codon to the termination codon (Table 1 and Fig. 1). The computer programs GeneScan (Burge and Karlin 1997) and FGENESH (Salamov and Solovyev 2000) also predicted a partial gene in reverse orientation to the MnSOD gene, encompassing bases 0 to 2,557 (Fig. 1). BLAST searches of GenBank revealed highly significant homology of this DNA segment to two S-locus receptor-like kinases in rice (GenBank accessions AF403128 and AF403126).

The MnSOD gene transcription start site was predicted at position 6308; 24 nucleotides prior to the start codon, using the TRANSFAC® program (Matys et al. 2003). Alignment of the genomic sequence with wheat MnSOD cDNA sequence (GenBank accessions AF092524, U72212, and U73172) and splice-junction identification with the program Splice View (Rogozin and Milanesi 1997) revealed the genomic MnSOD gene was composed of six exons and five introns (Table 1 and Fig. 1). The deduced amino acid sequence had the most similarity to that of GenBank accession U72212 with 99.1 % identity- only two amino acid differences out of 231. All introns in this MnSOD gene start with nucleotides GT at the 5' boundary and end with nucleotides AG at the 3' boundary. The intron splicing follows the GU-AG rule for the consensus sequences of splice junction sites (Breathnach and Chambon 1981; Buchanan et al. 2000), observed in the splice junctions of the rubber tree (Miao and Gaynor 1993) and human (Wan et al. 1994)

MnSOD genes.

The CpG Island Searcher program (Takai and Jones 2002, 2003) found there is a CpG island, an indicator of promoters in plants and animals, encompassing the 5' flanking region, exon1, and intron 1 of the MnSOD gene (Fig. 1). The CpG island, predicted by program parameters set to % GC=55, Obs CpG/Exp CpG=0.65, Length =500, and Distance=100 (Takai and Jones 2002, 2003), was 1000 bp in length, ranging from position 5963 to 6962 bp. PlantCARE software (Lescot et al. 2002) predicted cis-acting regulatory elements in the 5' flanking region of the MnSOD gene. PlantCARE indicated a TATA box at -380 and CAAT box at -369, -343, and -296; however, TATA boxes are generally located 25 to 30 bp upstream from transcription start sites and CAAT boxes about 50 bp upstream from the TATA box (Prescott 1988). Furthermore, program TSSP for predicting plant promoters predicted a TATA box in MnSOD genes in *Arabidopsis*, *Hevea*, *Pisum*, and *Oryza* (GenBank accessions GI 12408720, GI 348136, GI 945043, and GI 6440979, respectively), but not in the wheat MnSOD gene reported here.

The PlantCARE program predicted many cis-acting regulatory elements previously reported in monocotyledonous plants, including AAGAA-motif (-320; function unknown), AE-box (-246; part of a module for light response), CE1 (-442; cis-acting element associated with ABA responsive element, involved in abscisic acid responsiveness), G-box (-181; cis-acting regulatory element involved in light responsiveness), GC-motif (-77; enhancer-like element involved in anoxic specific inducibility), GC-repeat (-147 and -79; promoter), GCN4-motif (-210; cis-regulatory element involved in endosperm expression), I-box (-321; part of a light responsive element), P-box (-166; gibberellin-responsive element), RY-element (-469;

gibberellin-responsive element), and plant AP-2-like (-138 and -133; transcription factor).

The 3' UTR of the gene we sequenced proved to have a 17 nucleotide gap, compared to two of the three cDNA sequences in GenBank, but the third cDNA sequence in GenBank also had the gap (Fig. 2). A second gap of three nucleotides also was found in accession U72212 (Fig. 2). The significance of these gaps is unknown, but may impact RNA stability. Two polyadenylation signals (nucleotide position 11156: GTAATAAAAG and 11278: CTAATAAACT) were identified by HCPolya software (Milanesi et al. 1996).

Transposon insertion in intron 5

BLAST nucleotide-nucleotide searches revealed a 127 bp transposable element embedded in intron 5 of the MnSOD gene. TREP (the **T**riticeae **R**epeat Sequence Database, <http://wheat.pw.usda.gov/ITMI/Repeats/blastrepeats3.html>) indicated this sequence occurs numerous times throughout *Triticeae* genomes. This foldback element is classified as a form of miniature inverted-repeat transposable element (MITE; Fig. 3). The element is terminated in a CTCCTCC motif flanked by a TA target site duplication characteristic of MITEs. The transposable element is very similar to TREP1040 foldback element, MITE, stowaway Icarus_AV836460 and TREP1050 foldback element, MITE, stowaway Icarus_BE517313 118.

PCR with wheat genomic DNA from wheat accession 442 (Storlie et al. 1998) and cultivar 'El Tan' was performed with primers flanking intron 5 to investigate whether the *Stowaway* -like transposable element was inserted in all MnSOD genes. PCR with the genomic DNA templates generated two distinct bands of about 900 and 1,000 bp (Fig. 4a) in both wheat accessions. The

two bands were cloned from ‘El Tan’ and sequenced. As expected, the larger fragment contained the *Stowaway*-like element while the smaller fragment did not. The sequence of the PCR amplified *Stowaway*- like element we examined was 98 % identical (2 bp difference out of 127 bp) to the element found in the BAC clone, suggesting that multiple variants of this element occur in the various MnSOD genes.

We then examined mRNA transcripts from the same wheat lines to investigate whether the insertion sequence was involved in transcription. We examined transcripts from leaves and roots, and because we have shown previously that wheat MnSOD is upregulated in leaves in line 442 in response to cold temperature (Baek and Skinner, 2003), we also examined mRNA transcripts from leaves of line 442 that had been exposed to cold temperature. The amplifications of transcripts from roots and leaves grown at 20°C or at MnSOD-inducing conditions all resulted in a single fragment of about 100 bp., indicating the target consisted of only exons without the intron or the *Stowaway*-like element (Fig. 4b). This result suggested either there was no splicing variation caused by the inserted transposable element, or genes with the *Stowaway*-like element were not transcribed.

MnSOD Exons and Intron Positions in Plant Species

The wheat MnSOD gene sequence was compared with the DNA sequences of MnSOD genes in *Arabidopsis thaliana* (GenBank accession GI12408720), *Hevea brasiliensis* (GenBank accession GI348136), *Pisum sativum* (GenBank accession GI945043), and *Oryza sativa* (GenBank accession GI6440979), to establish the relative arrangement of the exons and introns (Table 1). The MnSOD genes from each of the plant species contained the same number of exons and

introns. However, the total length from the start codon to the termination codon ranged from 1567 bp in *Arabidopsis* to 4770 in wheat (Table 1). Virtually all of the size variation resulted from differences in intron length; the encoded protein products were either 231 aa or 233 aa (Table 1).

Two of the wheat MnSOD cDNA sequences in GenBank (accessions U72212 and U73172) encoded 231 aa products. However, GenBank accession AF 092524 encoded a 225 aa protein, due to a transit peptide six amino acids shorter (21 aa) than the other accessions (27 aa). Also, five wheat MnSOD cDNA clones from our laboratory all encoded 225 aa proteins that had the shorter transit peptide, identical to GenBank accession AF 092524. One cDNA from the TIGR database (accession TC207511) was reported to have a 39 aa transit peptide resulting in a 243 aa protein product. These differences in transit peptide length may indicate some of the MnSOD protein products are targeted to locations other than the mitochondria, such as different types of peroxisomes (Del Río et al. 2002).

Discussion

The region containing the promoter and the first exon of the wheat MnSOD gene was predicted to have a CpG island (Fig. 1). CpG islands are short stretches of DNA with higher frequency of the –CG– sequence. CpG islands are involved in gene regulation (Rombauts et al. 2003), and are found in association with about 60 % of human genes (Ioshikhes and Zhang 2000) and about 80 % of *Arabidopsis thaliana* and rice genes (Ashikawa 2001). The CpG island we identified in wheat was 1 kb in length, and ranged from position -367 to + 632 bp based on the start codon of the MnSOD gene. Gene-associated CpG islands usually are not methylated and usually are

linked to transcriptionally active DNA (Ashikawa 2001), in contrast to general methylation of CpG-rich regions (Rombauts et al. 2003). In noncoding regions, many cytosine residues in the CpG-rich regions are methylated on opposite DNA strands in the palindromic sequence of CpG. (Wolffe and Matzke 1999). DNA methylation is a powerful mechanism for the suppression of gene activity and there is a reciprocal relationship between the density of methylated cytosine residues and the transcriptional activity of a gene (Jeddeloh et al. 1998; Wolffe and Matzke 1999). Hence, the presence of the CpG island in the wheat MnSOD gene is suggestive of a methylation-dependent regulatory mechanism.

Increased transcript levels of wheat MnSOD genes in response to low temperature have been reported (Wu et al. 1999; Baek and Skinner 2003). The transcript increase may be a result of the CE1, ABA cis-acting regulatory element at position -442, an element shown to be responsive to accumulated ABA triggered by low temperature treatment (Lalk and Doerffling 1985; Doucette and Pritchard 1993). The AP2-like elements located at positions -138 and -33 also may be involved in cold-stress response. Similar elements are known to be involved in response to drought, high salinity, and cold temperature in other plant systems including the monocot *Zea mays* (Qin et al. 2004). Intense light resulted in increased expression of maize MnSOD (White and Scandalios, 1988). The wheat MnSOD gene had three light-responsive cis-acting regulatory elements, AE-box (-246), G-box (-181), and I-box (-321; part of a light responsive element), suggesting the wheat gene may also be light-responsive. Many of these cis-acting regulatory elements occur in similar locations in wheat and rice genomic MnSOD genes, such as AAGAA-motif (wheat -320, rice -296), G-box (wheat-181, rice -200), GC-motif (wheat -77, rice -87), GC-repeat (wheat -79, rice -89), and GCN4-motif (wheat -210, rice -196), suggesting

these elements may be evolutionarily conserved in monocots. However, *in silico* prediction of promoters and regulatory motifs is not straightforward (Rombauts et al 2003); further research is needed to determine whether the features identified by PlantCARE software are functional.

The feature prediction software suggested the wheat MnSOD gene we sequenced may not have a typical TATA box. A TATA box is also missing from the human MnSOD gene (Wan et al. 1994), the majority of plant genes involved in photosynthesis (Nakamura et al., 2002), and many housekeeping genes (Dyanan, 1986; Chye et al., 1992). However, because the feature prediction software is inexact, further study of this gene may reveal a TATA box or related features of similar function, as seen with genes involved in photosynthesis (Nakamura et al., 2002).

The 3' UTR of the wheat MnSOD gene appeared to have two polyadenylation signals, which theoretically could produce RNAs with different 3'-terminal poly (A) regions, resulting in alternative 3'-UTRs. The mechanism of differential poly (A) site selection probably results from regulating the competition between the splicing apparatus and the poly-A-site cleavage apparatus. The presence of multiple poly (A) signals has been demonstrated to result in more efficient and proper use of the terminal poly (A) signal (Russnak and Ganem 1990). There were no obvious differences in the size of MnSOD transcripts on a Northern blot with total RNA extracted from leaves (data not shown), therefore, it seems the first poly(A) site is not used for terminating wheat mRNA, but may be used for assisting the second poly (A) site for efficient and proper use.

A *Stowaway*-like transposable element was identified in intron 5 of the MnSOD gene we

sequenced. PCR analysis showed that not all copies of intron 5 had this element (Fig. 4a), and also that at least one sequence of intron 5 that lacked the element differed from the sequence that contained the element by two base pairs at the insertion site (not shown). This difference probably indicated a specific recognition site was necessary for successful insertion. Although the *Stowaway*-like transposable element was inserted in intron 5 of some copies of the MnSOD gene, there were no size variants observed in the mRNA transcripts, indicating either the insertion did not hinder splicing, or the insertion silenced the gene. A second *Stowaway*-like element was cloned from a different cultivar ('El Tan') and was found to differ in two nucleotides out of 127 from the element in the gene we sequenced. This difference may indicate that the insertion of this element was an ancient event, and changes have occurred in the inserted elements over time. Joining of the exon sequences from the gene we sequenced resulted in a complete, in-frame, coding sequence (not shown), indicating no indels or nonsense mutations have occurred, suggesting this particular gene is actively transcribed while the *Stowaway*-like element is maintained in intron 5.

All plant MnSOD genes investigated to date contained the same number and order of exons and introns. Although the size of the MnSOD genes varied from 1.5 kb in *Arabidopsis* to 4.8 kb in wheat, the encoded proteins differed by no more than two amino acids (Table 1). Considering that MnSOD is essential for protecting mitochondria from oxidative stress and maintaining specific concentrations of superoxide radicals, it is not surprising that the exons are highly conserved.

Upregulation of MnSOD transcription has been reported in response to numerous environmental cues. We found there are numerous regulatory elements in the promoter region of wheat MnSOD, including elements responsive to environmental stimuli and to abscisic acid. These findings indicate the wheat MnSOD gene may itself be responsive to some environmental variations and also may be part of the complex ABA regulatory network.

References

- Allen, R.D. 1995. Dissection of oxidative stress tolerance using transgenic plants. *Plant Physiol.* **107**: 1049-1054.
- Altschul, S.F., Madden T.L., Schaffer, A.A., Zhang Z., Miller W., Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389-3402.
- Ashikawa, I. 2001. Gene-associated CpG islands in plants as revealed by analyses of genomic sequences. *Plant J.* **26**: 617-625.
- Baek, K.-H., and Skinner, D.Z. 2003. Alteration of antioxidant enzyme gene expression during cold acclimation of near-isogenic wheat lines. *Plant Sci.* **165**: 1221-1227.
- Basu, U., Good, A.G., and Taylor, G.J. 2001. Transgenic *Brassica napus* plants overexpressing aluminum-induced mitochondrial manganese superoxide dismutase cDNA are resistant to aluminum. *Plant, Cell, and Environ.* **24**: 1269-1278.
- Bowler, C., Alliotte T., Van Den Bulcke, M., Bauw, G., Vandekerckhove, J., Van Montagu, M., and D. Inze. 1989. A plant manganese superoxide dismutase is efficiently imported and correctly processed by yeast mitochondria. *Proc. Natl. Acad. Sci. USA* **86**: 3237-3241.
- Bowler, C., Slooten, L., Vandenbranden, S., De Rycke, R., Botterman, J., Sybesma, C., Van Montague, M., and Inze, D. 1991. Manganese superoxide dismutase can reduce cellular damage mediated by oxygen radicals in transgenic plants. *EMBO J.* **10**: 1723-1732.
- Bowler, C., Montagu, M.V., and Inze, D. 1992. Superoxide dismutase and stress tolerance. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**: 83-116.
- Breathnach, R., and Chambon, P. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**: 349-383.

- Breusegem, F, Van, Slooten, L., Stassart, J.M., Botterman, J., Moens, T., Van Montagu, M., and Inze, D. 1999. Effects of overproduction of tobacco MnSOD in maize chloroplasts on foliar tolerance to cold and oxidative stress. *J. Exp. Bot.* **50**: 71-78.
- Buchanan, B., Gruissem, W., and Jones R.L. 2000. *Biochemistry & molecular biology of plants*. American Society of Plant Physiologists, Rockville, Maryland, p 302.
- Burge, C., and Karlin, S. 1997. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**: 78-94.
- Carlioz ,A., and Touati, D. 1986. Isolation of superoxide dismutase mutants in *Escherichia coli*: is superoxide dismutase necessary for aerobic life? *EMBO J.* **5**: 623-630.
- Chye, M.-L., C.-T. Tan, and N.-H. Chua. 1992. Three genes encode 3-hydroxy-3-methylglutaryl-coenzyme A reductase in *Hevea brasiliensis*: hmg1 and hmg3 are differentially expressed. *Plant Mol. Biol.* **19**:473-484.
- Clare D.A., Rabinowitch, H.D., and Fridovich, I. 1984. Superoxide dismutase and chilling injury in *Chlorella ellipsoidea*. *Arch. Biochem. Biophys.* **231**: 158-168.
- Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* **16**: 10881-10890.
- Del Río, L.A., Corpas, F.J., Sandalio, L.M., Palma, J.M., Gomez, M., and Barroso, J.B. 2002. Reactive oxygen species, antioxidant systems and nitric oxide in peroxisomes. *J. Exp. Bot.* **53**: 1255-1272.
- Doucette, M.S. and Pritchard, M.K. 1993. ABA involvement in low temperature sweetening of potatoes. *Acta. Hort.* **343**:293-294.
- Dynan, W.S. 1986. Promoters for housekeeping genes. *Trends Genet.* **2**: 196-197.
- Fink, R.C., and Scandalios, J.G. 2002. Molecular evolution and structure- function relationships

- of the superoxide dismutase gene families in angiosperms and their relationship to other eukaryotic and prokaryotic superoxide dismutases. *Arch. Biochem. Biophys.* **399**: 19-36.
- Fridovich, I. 1991. Molecular oxygen; Friend and foe. In *Active oxygen species/oxidative stress and plant metabolism*. Edited by E. J. Pell and K. L. Steffen. American Society of Plant Physiologists, Rockville, MD. pp 1-5.
- Fucci, L., Oliver, C.N., Coon, M.J., and Stadtman, E.R. 1983. Inactivation of metabolic enzymes by mixed-function oxidation reaction: possible implication in protein turnover and ageing. *Proc. Natl. Acad. Sci. USA* **80**: 1521-1525.
- Halliwell, B., and Gutteridge, J.M.C. 1999. *Free radicals in biology and medicine*. 3rd ed. Oxford University Press, New York. pp 262-282.
- Huang, X., and Madan, A. 1999. CAP3: A DNA sequence assembly program. *Genome Res.* **9**: 868-877.
- Ioshikhes, I.P., and Zhang, M.Q. 2000. Large-scale human promoter mapping using CpG islands. *Nat. Gene.* **26**: 61-63.
- Jeddeloh, J.A., Bender, J., and Richards, E.J. 1998. The DNA methylation locus DDM1 is required for maintenance of gene silencing in Arabidopsis. *Genes. Dev.* **12**: 1714-1725.
- Kendall, E.J., and McKersie, B.D. 1989. Free radical and freezing injury to cell membranes of winter wheat. *Physiol. Plant.* **76**: 86-94.
- Lalk, I., and Doerffling, K. 1985. Hardening, abscisic acid, proline, and freezing resistance in two winter wheat varieties. *Physiol. Plant.* **63**: 287-292.
- Lescot, M., Déhais, P., Moreau, Y., De Moor, B., Rouzé, P., and Rombauts, S. 2002. PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* **30**:325-327.

- Ling, P. and Chen, X. 2006. Construction of a hexaploid wheat (*Triticum aestivum* L.) bacterial artificial chromosome library and identification of candidate BAC clones containing stripe rust resistance gene markers. *Genome in press*.
- Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. 1999. Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *J. Mol. Biol.* **288**, 911-940.
- Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., Kloos, D.U., Land, S., Lewicki-Potapov, B., Michael, H., Munch, R., Reuter, I., Rotert, S., Saxel, H., Scheer, M., Thiele, S., and Wingender, E. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **31**:374-378.
- McKersie, B.D., Chen, Y.R., De Beus, M., Bowler, S.R., Inze, D., D'alluin, K., and Botterman, J. 1993. Superoxide dismutase enhances tolerance of freezing stress in transgenic alfalfa (*Medicago sativa* L.). *Plant Physiol.* **103**: 1155-1163.
- Miao, Z., and Gaynor, J.J. 1993. Molecular cloning, characterization and expression of Mn-superoxide dismutase from the rubber tree (*Hevea brasiliensis*). *Plant Mol. Biol.* **23**: 267-77.
- Milanesi, L., Muselli M., and Arrigo, P. 1996. Hamming-clustering method for signals prediction in 5' and 3' regions of eukaryotic genes. *Comput. Applic. Biosci.* **12**: 399-404.
- Møller, I.M. 2001. Plant mitochondria and oxidative stress: Electron transport, NADPH turnover, and metabolism of reactive oxygen species. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**: 561-591.
- Monk, L.S., Fagerstedt, K.V., and Crawford, R.M.M. 1989. Oxygen toxicity and superoxide

- dismutase as an antioxidant in physiological stress. *Physiol. Plant.* **76**: 456-459.
- Nakamura, M., T. Tsunoda, and J. Obokata. 2002. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. *Plant J.* **29**: 1-10.
- Prescott, D.M. 1988. *Cells*. Jones and Bartlett Publishers, Inc., Boston London. pp 314-315.
- Qin, F., Sakuma, Y., Li, J., Liu, Q., Li, Y.Q., Shinozaki, K., and Yamaguchi-Shinozaki, K. 2004. Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant Cell Physiol.* **45**: 1042-1052.
- Rogozin, I.B., and Milanesi, L. 1997. Analysis of donor splice signals in different organisms. *J. Mol. Evol.* **45**: 50-59.
- Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouzé, P., and Van de Peer, Y. 2003. Computational approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant Physiol.* **132**: 1162-1176.
- Russnak, R., and Ganem, D. 1990. Sequences 5' to the polyadenylation signal mediate differential poly(A) site use in hepatitis B viruses. *Gene. Dev.* **4**: 764-776.
- Salamov, A., and Solovyev, V. 2000. Ab initio gene finding in *Drosophila* genomic DNA. *Genome Res.* **10**: 516-522.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning: A laboratory manual*. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Sato, I., Zu, J., Nishikawa, S., and Kashimura, N. 1993. Depolymerization of hyaluronic acid by D-fructose 6-phosphate. *Biosci. Biotechnol. Biochem.* **57**: 2005-2009.
- Storlie, E.W., Allan, R.E., and Walker-Simmons, M.K. 1998. Effect of the Vrn1-Fr1 interval on cold hardiness levels in near isogenic wheat lines. *Crop Sci.* **38**: 483-488.

- Streller, S., Kromer, S., and Wingsle, G. 1994. Isolation and purification of mitochondrial Mn-superoxide dismutase from the gymnosperm *Pinus sylvestris* L. *Plant Cell Physiol.* **35**: 859-867.
- Takai, D., and Jones, P.A. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. USA* **99**: 3740-3745.
- Takai, D., and Jones, P.A. 2003. The CpG island searcher: A new WWW resource. *In Silico Biology* **3**: 235-240.
- Tatusova, T.A., and Madden T.L. 1999. BLAST 2 sequences - a new tool for comparing protein and nucleotide sequences. *FEMS Microbiol Lett.* **174**: 247-250.
- Wan, X.S., Devalaraja, M.N., and St. Clair, D.K. 1994. Molecular structure and organization of the human manganese superoxide gene. *DNA Cell Biol.* **13**: 1127-1136.
- Wellings, C.R., Singh, R.O., McIntosh, R.A., and Pretorius, Z.A. 2004. The development and application of near isogenic lines for the stripe (yellow) rust pathosystem. *Proceedings of the 11th International Cereal Rusts and Powdery Mildew Conference, Norwich, England, 22-27 August 2004. Abstract A1.39, Cereal Rusts and Powdery Mildews Bulletin.*
- White, J.A., and Scandalios, G. 1988. Isolation and characterization of a cDNA for mitochondrial manganese superoxidedismutase (SOD-3) of maize and its relation to other manganese superoxide dismutase. *Biochim. Biophys. Acta* **951**: 61-70.
- Wolffe, A.P., and Matzke M.A. 1999. Epigenetics: Regulation through repression. *Science* **286**: 481-486.
- Wu, G., Shortt, B.J., Lawrence, E.B., Fitzsimmons, J.L.K.C., Levine, E.B., Raskin, I., and Shah, D.M. 1997. Activation of host defense mechanisms by elevated production of H₂O₂ in transgenic plants. *Plant Physiol.* **115**: 427-435.

- Wu, G., Wilen, R.W., Robertson, A.J., and Gusta, L.V. 1999. Isolation, chromosome localization, and differential expression of mitochondrial manganese superoxide dismutase and chloroplastic copper/Zinc superoxide dismutase genes in wheat. *Plant Physiol.* **120**: 513-520.
- Yu, Q., Osborne, L.D., and Rengel, Z. 1999. Increased tolerance to Mn deficiency in transgenic tobacco overexpressing superoxide dismutase. *Ann. Bot.* **84**: 543-547.
- Zelko, I.N., Mariani, T.J., and Folz, R.J. 2002. Superoxide dismutase multigene family: A comparison of CuZn-SOD (SOD1), Mn-SOD (SOD2), and EC-SOD (SOD3) gene structures, evolution, and expression. *Free Radic. Biol. Med.* **33**: 337-349.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**: 3406-3415.

Figure Captions.

Figure 1. Map of an 11.3 kb sequence of wheat genomic DNA containing the MnSOD gene sequence and part of a kinase gene. The boxes depict the locations of exons. The locations of the subclones and a PCR product used in the construction of the entire sequence are also shown.

Figure 2. Alignment of 3' UTR sequences of wheat MnSOD cDNAs (GenBank accessions AF092524, U72212, and U73172) and the genomic MnSOD sequenced in this study (GenBank accession AY963808). Asterisks indicate consensus among all four sequences.

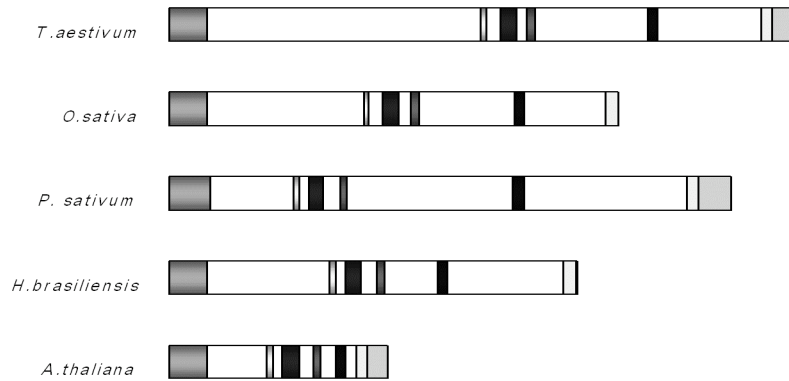
Figure 3. DNA sequence and predicted secondary structure of an apparent transposable element in intron 5 of a wheat MnSOD gene. The *Stowaway*-like transposable element was predicted to form a hair-pin structure with -67.5 kcal/mol free energy of dissociation.

Figure 4. Gel pictures of PCR (a) and RT-PCR (b) using forward primer 5'-TCCTTTGTTGGGAATTGATGTC in exon 5 and reverse primer 5'-TCACGCAAGCACTTTTTCATACTCT in exon 6. (a) PCR using DNA from wheat line 442 or 'El Tan' as template. The two arrows indicate the two PCR fragments, of about 1000 bp and 900 bp. Lane 1: 1 kb plus ladder from Invitrogen. Lane 2 and 3: PCR with 442 and 'El Tan' DNA with 52 °C annealing temperature, respectively. Lane 4 and 5: PCR products as in lanes 2 and 3 but 58 °C annealing temperature. (b) One-step RT-PCR using mRNA template or PCR using cDNA template from different growing conditions or tissues. Lanes 1 and 2: One-step RT-PCR using mRNA from wheat 442 leaves grown at 20 °C for 2 weeks (lane 1), or 2 °C for 4 weeks

following 2 weeks growth at 20 °C (lane2). Lane 3 and 4: PCR using total cDNA from leaves (lane 3) or roots (lane 4) of 442 plants grown at 20 °C for 2 weeks. Lane 5: 1 kb plus ladder from Invitrogen.

Table 1. Comparison of five different plant genomic MnSOD gene structure. The picture in the table was drawn to show the locations of exons and introns. All blanks are introns, and exon 6 is divided into two parts, indicating the termination codon.

	<i>A.thaliana</i>	<i>H.brasiliensis</i>	<i>P. sativum</i>	<i>O.sativa</i>	<i>T.aestivum</i>
Exon1 (from the translation start codon)	295	295	322	295	295
Intron 1	478	972	657	1241	2162
Exon 2	47	47	47	47	47
Intron 2	80	78	83	108	114
Exon 3	126	125	111	126	126
Intron 3	111	123	133	96	90
Exon 4	57	58	48	57	57
Intron 4	119	428	1325	763	899
Exon 5	78	78	81	78	78
Intron 5	83	909	1287	648	812
Exon 6 (to the termination codon)	93	99	93	93	93
Exon 6 (3' UTR)	165	22	256		187
Initiation codon to termination codon	696	702	702	696	696
No. of Amino Acids	231	233	233	231	231
Total Gene Size (from the start codon to the termination codon)	1567	3212	4187	3552	4770



Note: The journal version of this table removed the figure from the bottom and converted it to a separate figure.

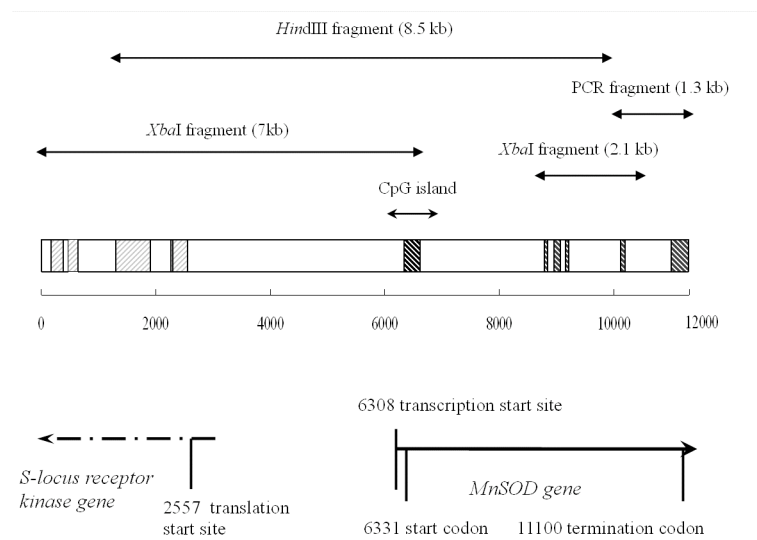


Fig 1.

```

U73172      TTTGTCTGATGAGTAGACAGACGGTACTTTGCTCGCTCTTGAGTTT-----
AY963808    TTTGTCTGATGAGTAGACAGACGGTACTTTGCTCGCTCTTGAGTTT-----
AF092524    TTTGTCTGATGAGTAGACAGACAGTACTTGGCTCGCTCTTGAGTTTCCCGCTGTTGGATC
U72212      TTTGTCCGA---GTAGACAGACAGTACTTTCGCTCGCTCTTGAGTTTCCCGCTGTTGGATC
          ***** **      ***** ***** *****

U73172      ---GTATATGCAGTAATAAAAAGTGGGCACTATTCCACTGCCGGGATCGTCTGTACATTGC
AY963808    ---GTATATGCAGTAATAAAAAGTGGGCACTATTCCACTGCCGGGATCGTCTGTACATTGC
AF092524    TTTGTATATGCAGTAATAAAAAGTGGGCACTATTCCACTGTCGGGACCTTCTGTACATTGC
U72212      TTTGTATATGCAGTAATAAAAAGTGGGCACTATTCCACTGTCGGGACCTTCTGTACATCGC
          ***** * ***** **

U73172      ATAGAGGTGG-ACTACTGTCTAATGCGCACTACAGTGCCTGTCCATTTTGGTACTTGTTA
AY963808    ATAGAGGTGG-ACTACTGTCTAATGCGCACTACAGTGCCTGTCCATTTTGGTACTTGTTA
AF092524    ATAGAGGTGACACTACTGTCTAAAGCGCACTA--GTGCATGTCCATTTTGGTACTTGTTA
U72212      ATAGAGGCGG-GGTAATGTCCAACGCATAGAACAGTGCCTGTCCATTTTGGTACTTTTTA
          ***** *      ** ***** ** *      *      ***** *****

U73172      CTGTGCCTGCTGTGGCTAATAAA-----
AY963808    CTGTGCCTGCTGTGGCTAATAAACTCGTTCGCTAGTGCTGTT
AF092524    CTGTGCCTGC-GTTGCTAATAAACTCATTCGCTAGTGCTGTT
U72212      CTGTGCCTGC-GTGACTAATAAACTCGTTCGCTAGTGCTGTT
          ***** **      *****

```

Fig. 2.

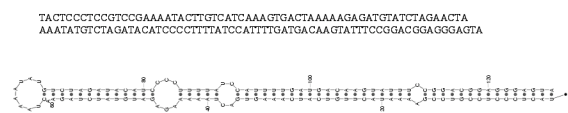


Fig 3.

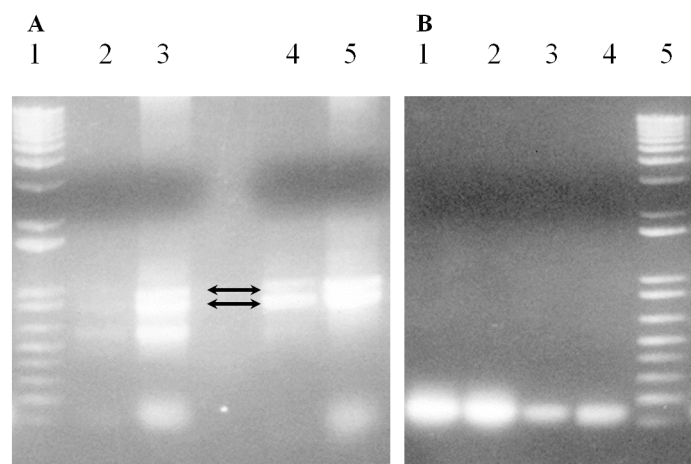


Fig 4.